DOCUMENT RESUME

TM 000 656

ED 052 250

AUTHOR	Bianchini, John C.; And Others
TITLE	The Berkeley Project: Race and Socio-Economic Status in the Selection Testing of Municipal Personnel.
PUB DATE	Oct 66
NOTE	30p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	*Aptitude Tests, Caucasians, Community Agencies (Public), *Competitive Selection, Government Employees, Item Analysis, Minority Groups, Negroes, Performance Specifications, *Personnel Selection, *Racial Differences, Scores, *Socioeconomic Status, Test Bias, Testing
IDENTIFIERS	Otis Quick Scoring Mental Ability Test, *The Berkeley Project

ABSTRACT

Racial factors in personnel selection procedures and the performance of racial and socio-economic groups on a testing instrument widely used in personnel selection and in public schools are discussed. The study questions whether individual test items are relatively easier for one racial group than for another, and whether there is a relationship between test performance and socio-economic status. A description of the instrument, subjects, method, and a careful analysis of the data is included. Considerable interest has been aroused by this study, perhaps because it is the first project of its kind to be undertaken. (TA)

ED052250

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

THE BERKELEY PROJECT: RACE AND SOCIO-ECONOMIC STATUS IN THE SELECTION TESTING OF MUNICIPAL PERSONNEL

TM 000 656

by

John C. Bianchini
Educational Testing Service

William F. Danielson
City of Berkeley

Robert W. Heath
Stanford University

Clinton A. Hilliard
City of Berkeley

October, 1966

THE "BERKELEY PROJECT"

Recently there has been extensive expression of concern about the use of written tests in personnel selection. In particular, the possibility that the use of such tests may invalidly prevent minority group members from obtaining jobs has been the subject of public discussion.

In this report the performance of racial and socio-economic groups on a testing instrument widely used in personnel selection and in public schools is described. The purpose of the study was fourfold:

(a) To determine to what degree test performance differed for members of different racial groups;

(b) To ascertain if individual test items were relatively easier for one racial group than for another and to analyze each such item;

(c) To determine the relationship between test performance and socio-economic status;

(d) To identify needs for further research.

The City of Berkeley has had a formal, positive personnel program with relation to minority groups since August, 1963. At that time the City Council of Berkeley requested the Personnel Department and the Personnel Board to undertake a continuing program to find means of improving personnel procedures, using any appropriate constructive research methods.

In December, 1963, the California Fair Employment Practice Commission revised its rules to permit California Employers to keep certain group records of racial origin of candidates and of employees for constructive research purposes. As soon as the California F.E.P.C. had given this permission, the City of Berkeley Personnel Department began to keep group records of the racial origin of candidates participating

in competitive examinations for city employment. The method used was that the monitor of the test would estimate the number of candidates by race who appeared at each phase of the selection process. The monitor estimated how many candidates were Caucasians, Negroes, Orientals, Spanish speaking/surname, or members of other ethnic groups. The monitor listed these tallies separately from the examination results. The monitor did not identify the racial origin of any individual candidate.

With six to seven months' cumulation of data of this kind, a pattern began to emerge. The pattern indicated that there were no significant differences in the performance of racial groups at any stage of the selection process except at the written test level. With regard to written tests, very significant differences occurred in the performance of the various racial groups.

In the examples below, statistics are shown of data collected during this period of time for candidates for Fireman, Laborer, Intermediate Typist Clerk, Electrician Helper, and Public Health Nurse.

EXAMINATION TITLE	C	N	Or	Sp.s/s	O	TOTAL
Fireman/6-16-64						
NUMBER TESTED	166	32	-	6	-	204
Passed Written	84	2	-	4	-	90
Physical Agility Test	75	2	-	4	-	81
Passed Physical Agility Test	47	1	-	3	-	51
Number Interviewed	43	1	-	3	-	47
Passed Interview	36	1	-	3	-	40
Public Health Nurse/1-8-64						
NUMBER TESTED	6	3	-	-	-	9
Passed Written	6	3	-	-	-	9
Interviewed	6	2	-	-	-	8
Passed Interview	5	2	-	-	-	7

EXAMINATION TITLE	C	N	Or	Sp.s/s	O	TOTAL
Electrician Helper 12-4, 5-63						
NUMBER TESTED	62	26	1	-	-	89
Passed Written	-	-	-	-	-	18
Interviewed	14	1	1	-	-	16
Passed Interview	10	1	1	-	-	12
Laborer/2-10, 11, 13-64						
NUMBER TESTED	13	224	-	6	5	248
Passed Written	-	-	-	-	-	55
Strength & Agility Test	5	35	-	5	4	49
Passed Strength & Agility Test	5	32	-	5	4	46
Interviewed	5	31	-	5	4	45
Passed Interview	5	29	-	5	4	43
Intermediate Typist Clerk 4-15-64						
NUMBER TESTED	14	12	2	-	-	28
Passed Written	7	3	1	-	-	11
Skills Test (Typing)	6	2	1	-	-	9
Passed Skills Test	3	2	1	-	-	6
Interviewed	6	2	1	-	-	9
Passed Interview	5	2	1	-	-	8
Passed Skills Test and Interview	3	2	1	-	-	6

In the examples above, the two largest numbers of candidates by racial group are Negro and Caucasian. The Negro candidates failed on the written tests at a substantially higher rate than the Caucasians attempting the same tests except for the one examination which required a substantial amount of education and experience (Public Health Nurse, where the performance of Caucasians and Negroes is about the same).

The City of Berkeley, as can be seen by the examples given, like most public personnel agencies, selects candidates by requiring that the candidates pass each of several successive stages in the selection process. First, the candidate must establish that he meets the minimum qualifications (e.g., education, experience, license, etc.) which are required to compete for a given position. If the candidate is admitted

to the examination, the first step in the selection process is normally the administration of a written test designed to measure the aptitudes and/or knowledges required for the position. Those candidates who pass this written test battery are invited to an interview. For certain positions, other tests may be used (e.g., typing test for Intermediate Typist Clerk; physical agility test for Fireman).

Therefore, a central problem was identified: The examination statistics showed that Negro candidates were failing at a much higher rate than Caucasian candidates on the written tests. The results also showed that among those candidates passing the written tests, there was no significant difference on a racial basis, among those qualifying and failing to qualify on the interview, skills tests and other selection procedures.

Having identified that there is a significantly larger portion of Negro candidates failing on written tests than Caucasian, the next question was "why?" Were the written tests discriminating, unintentionally, on a basis of racial origin or not? What factors were causing the differences in performance of racial groups?

It was immediately apparent that no further investigation could be made into these questions unless individual candidates could be identified by race and certain other factors regarding socio-economic status. The California F.E.P.C. regulations forbid asking such questions prior to employment. It was necessary, however, to have this information before any kind of study could be attempted which could determine what really happens when written tests are used in a competitive job situation.

Mr. Danielson, Director of Personnel, City of Berkeley, contacted Dr.

Robert Heath, then serving as Research Director for the Educational Testing Service, West Coast Office, Berkeley, regarding the feasibility of undertaking a research project. Dr. Heath advised that a research project was feasible. He offered to serve as a consultant. The Educational Testing Service made available data processing machine time to analyze the results of such a study. The assistance of Dr. Heath made the project feasible by providing the expert psychological consulting and the data analysis services required to analyze test results.

Director of Personnel, Danielson, then approached the City Manager of Berkeley, Mr. John D. Phillips, to obtain consent to approach the Fair Employment Practice Commission for an exemption to conduct the research study. Upon being informed of the objectives of the study and the possible use of the results to enable a more meaningful positive personnel program to be carried out, Mr. Phillips approved the project.

Next, meetings were held with staff members of the Fair Employment Practice Commission. Mrs. Aileen Hernandez, then Assistant Director, and Mr. Hugh Taylor, Consultant, met with Dr. Heath and Mr. Danielson to review the proposed project and the exemption which would be required. Encouragement was given by F.E.P.C. staff members to proceed with the project because of the need for research in employment testing.

The Personnel Board of the City of Berkeley and the liaison member from the Council to the Personnel Board, Councilman Wilmont Sweeney, were consulted about the advisability of undertaking such a research project. Positive encouragement was provided by the Board and Councilman Sweeney. The Director of Personnel was advised to request the

F.E.P.C. exemption which would enable the "Berkeley Project" to proceed.

In December, 1964, the Director of Personnel met with the California Fair Employment Practice Commission to request a one-year exemption to obtain racial origin data on individual candidates. Permission was granted unanimously by the Fair Employment Practice Commission for the "Berkeley Project." The F.E.P.C. gave positive encouragement and stated that this was the first such request made by any California employer to undertake a research study for the purpose of a positive research program. At the same December, 1964, meeting of the Fair Employment Practice Commission, the Technical Advisory Committee on Testing (TACT) was created. A principal, initial purpose of the TACT Committee was to advise Dr. Heath and Mr. Danielson with regard to the "Berkeley Project" and to serve as advisors to the Fair Employment Practice Commission in the matter of employment testing. Dr. Heath and Mr. Danielson were named as original members of TACT.

In May, 1965, the Fair Employment Practice Commission approved the procedures to be used for the collection of data in the "Berkeley Project." At the same F.E.P.C. meeting, an exemption similar to that granted to the City of Berkeley in December, 1964, was granted to the City of Fresno. Mr. Wayne Higbee, then Personnel Director for the City of Fresno, had requested permission to join the City of Berkeley in this research project. As the City of Fresno has large numbers of Spanish speaking/surname candidates (which the City of Berkeley does not have) the F.E.P.C. granted this additional exemption.

A special optically-scanned test answer sheet was devised which was used by both the City of Berkeley and City of Fresno. The Otis Quick

Scoring Mental Ability Test, Gamma Test: Form Em, was selected to use in the research project. This instrument was selected for a number of reasons. It is widely used throughout the United States by many employers, both public and private, for screening purposes for many kinds and levels of jobs. The thirty minutes required for the test made it possible to include on an advisory basis with all examinations to be given by the City of Berkeley and by the City of Fresno.

The collection of data began in May, 1965. Data collection was stopped as of April 30, 1966, the end of the one-year exemption period.

Dr. John Bianchini of Educational Testing Service has directed the analysis of data collected.

The "Berkeley Project" has aroused great interest throughout the United States although it has been a limited research project because it is the first of its kind to be undertaken. The "Motorola Case" in Illinois has added interest to the results to be obtained from the "Berkeley Project" study.

In the 1965 session of the California Legislature, there were many bills introduced to regulate procedures to be used by public employers. Almost all of the bills presented to the Legislature were intended to restrict the use of the oral interview in the belief that it is at the interview that racial discrimination occurs in the public service. For example, bills were introduced which would have eliminated the use of the interview in public service and, therefore, would have made the written test the entire score, for ranking purposes, of candidates. Another bill would have restricted the weight of the interview to 10%

and, therefore, increase the weight of the written test to 90%.

Since December, 1963, when Berkeley first began to keep records of racial origin of candidates, many other public agencies of California also have done so. Twenty major counties in California and the California State Personnel Board have kept examination racial origin records through the present date. The results of these records were reported by Harold S. Rosen, Personnel Director, Santa Clara County, in his article, "Equal Opportunity Under the Merit System," in the July, 1966, Public Personnel Review. Mr. Rosen reported that among 40,000 applicants who had been tested, 42% of all Caucasian applicants had passed all written tests while 24% of all Negro applicants passed the tests. He also reported significant differences in the kinds of jobs for which application was made by the various racial groups. For example, 12% of Caucasian candidates applied for custodial or laboring jobs in the twenty counties and 53% of the Negro applicants applied for such jobs. The results obtained by the counties, by the State of California, and by other cities keeping such records correspond very closely to those obtained by the City of Berkeley: Among racial groups there are significant differences on the written test performance.

Procedure

Each candidate admitted to the testing room was asked to present an admission card which had been mailed to him. The test monitor verified the candidate's name and recorded the racial origin of each candidate by means of a coded symbol. At no time during this identification procedure were the candidates made aware that racial identification was taking place. The racial origin information recorded on the

admission list was later transcribed to the special questionnaire-answer sheet.

To preserve the anonymity of candidates until cutting scores on tests were set, each candidate was assigned an identification number prior to administration of the tests. This number was used to identify the test papers. The identification forms bearing the candidate's names were placed in a sealed envelope. After cutting scores were set, the envelope containing the identification forms was opened. At this point, the name, identification number and racial origin of each candidate taking the written test were matched. It should be stressed that this identification occurred after cutting scores had been set. Knowledge of the name and racial origin of each candidate did not affect the evaluation of his test score.

The candidates were instructed not to write their names on any of the test papers and were informed that the Otis test and special answer sheet were part of a research project which involved many of the employment examinations conducted by the City of Berkeley. Further, it was explained that the questionnaire section of the special answer sheet would not affect the candidates' chances for success in the selection process.

After the preliminary instructions were given to each group of candidates, the special questionnaire-answer sheet was distributed. To obtain the information from each candidate as accurately and completely as possible, the monitor read and explained the questionnaire section. Each examinee provided the following biographical information: candidate's education, present (or last) occupation, age, sex, language

spoken in the home, and mother's education. Each candidate also rated his father's occupation using the Social Class Identification Scale (SCI).[1] After the questionnaire section was completed by each candidate, the Otis test booklet was distributed.

Following the administration of the test, the scores were recorded on the Personnel Department test score form. The last three sections on the special answer sheet were then completed by the chief examiner. These asked the candidate's racial origin, the code number of the job for which the test was given and the socio-economic status rating of each candidate's present occupation. The racial origin coding was transcribed from the examination admission list.

If in fact, traditional intelligence tests are more difficult for certain racial and socio-economic groups, it should be possible to obtain information which makes apparent those factors relevant to the performance of various groups. Accordingly, an attempt was made to compare the responses of the different groups on a set of test items and to examine individual items for clues as to the nature of possible differences.

A comparison of item responses for different groups generates two hypotheses of interest. First, we would want to know if there are mean differences across groups, and second, if there are some items which are relatively easier for one group than for another. In other words, a method is needed which provides an estimate of the interaction of items with groups, since a significant interaction indicates that

---

(1) Heath, R.W. 'The development of a measure of Social Class Identification." Purdue Stu. High Educ., 1958, 87, 45 pp.

certain items are relatively easier for one group than for another.

Thus, the analysis of variance design for two-factor experiments with repeated measures on one factor is appropriate, since the different groups performed on the same set of items. It should be added that in this instance, both groups and items are treated as fixed effects and only samples are random within each group.

The design has the following schematic representation:

		Items			
		1	2	3	56
White	Sample 1				
	Sample 2				
Negro	Sample 1				
	Sample 2				

The assumptions for this model are (1) that the item variances are equal, and (2) that the inter-item covariances are equal. In order to satisfy the first assumption, the item p-values (proportion of subject in each sample passing a given item), were transformed to arcsin values, $\phi = 2 \arcsin \sqrt{p}$. The use of this transformation assures that the variance for the sampling distribution for any item regardless of the difficulty is $1/N$ and is dependent only on the size of the sample. The second assumption is assumed to be sufficiently satisfied.

Each sample consisted of 194 applicants selected at random in the following manner. The Negro applicants were assigned alternately to one of two samples. From the white group, a subgroup was chosen consisting of those applicants who had applied for occupations for which Negroes had applied also. From this subgroup, two samples were selected by drawing every fourth candidate. Mexican-American applicants were

not included in this analysis because of the relatively small number of such candidates.

Since an individual's education, occupation, and Social Class Identification score tend to reflect somewhat different dimensions of the socio-economic structure of our culture, it appeared desirable to develop a composite quantitative index of socio-economic status to obtain information about the relationship between test performance and socio-economic characteristics. Thus, both applicant's and mother's education were expressed on a 17 point scale, corresponding to a zero for kindergarten, a one for first grade and so on, up to a 17 for college work beyond the Baccalaureate degree. The Index of Social Class Identification scores was expressed on a 12 point scale. From these measures, a single socio-economic status index (SES) was derived, giving equal effective weights to each of the four components, viz., applicant's education, mother's education, applicant's Social Class Identification score based upon father's occupation, and examiner's rating of the applicant's occupation in terms of the Social Class Identification Scale. The SES measure was expressed on a scale having a mean of 50 and a standard deviation of 10.

To obtain information about the degree of relationship between differences in test performance and differences in the socio-economic status of the various racial groups, the analysis of covariance technique was chosen for analyzing the data. This method of analysis makes it possible to evaluate differences in test performance when actual differences in socio-economic status among the racial groups are held constant.

Results

The descriptive data which follow describe certain characteristics of the various racial groups which are relevant to the interpretation of the data. The data in Table 1 show the number of applicants tested for various civil service occupations and the racial origin of the applicants.

For a further description by racial group for those occupations having a sufficient number of applicants, the means, standard deviations and N's are reported in Table 2. From these data, the following trends should be noted. First, the ratio of the white applicants to Negro applicants is approximately 5:1 for the "higher status" occupations, i.e., Patrolman, Fireman and Recreation Leader, while for "lower status" occupations, i.e., Gardener, Laborer and Waste Collector, the ratio of white applicants to Negro applicants is approximately 1:2. Second, applicants for Patrolman, Fireman and Recreation Leader scored substantially higher on the OTIS than applicants for Gardener, Laborer and Waste Collector. Third, white applicants performed better on the OTIS than Negroes or Mexican Americans.

The frequency distribution of the OTIS scores and SES scores for the racial groups are shown in figures 1 and 2 respectively. From these graphs it may be seen that although the three distributions overlap, the mean for the white distribution is considerably higher than the average score of either the Negroes or Mexican-Americans. These data imply that Negroes and Mexican-Americans as groups have similar socio-economic characteristics and their test performance is also similar. However, the socio-economic characteristics of the white applicants are

TABLE 1

Frequencies of Applicants by Race for the Various Occupations

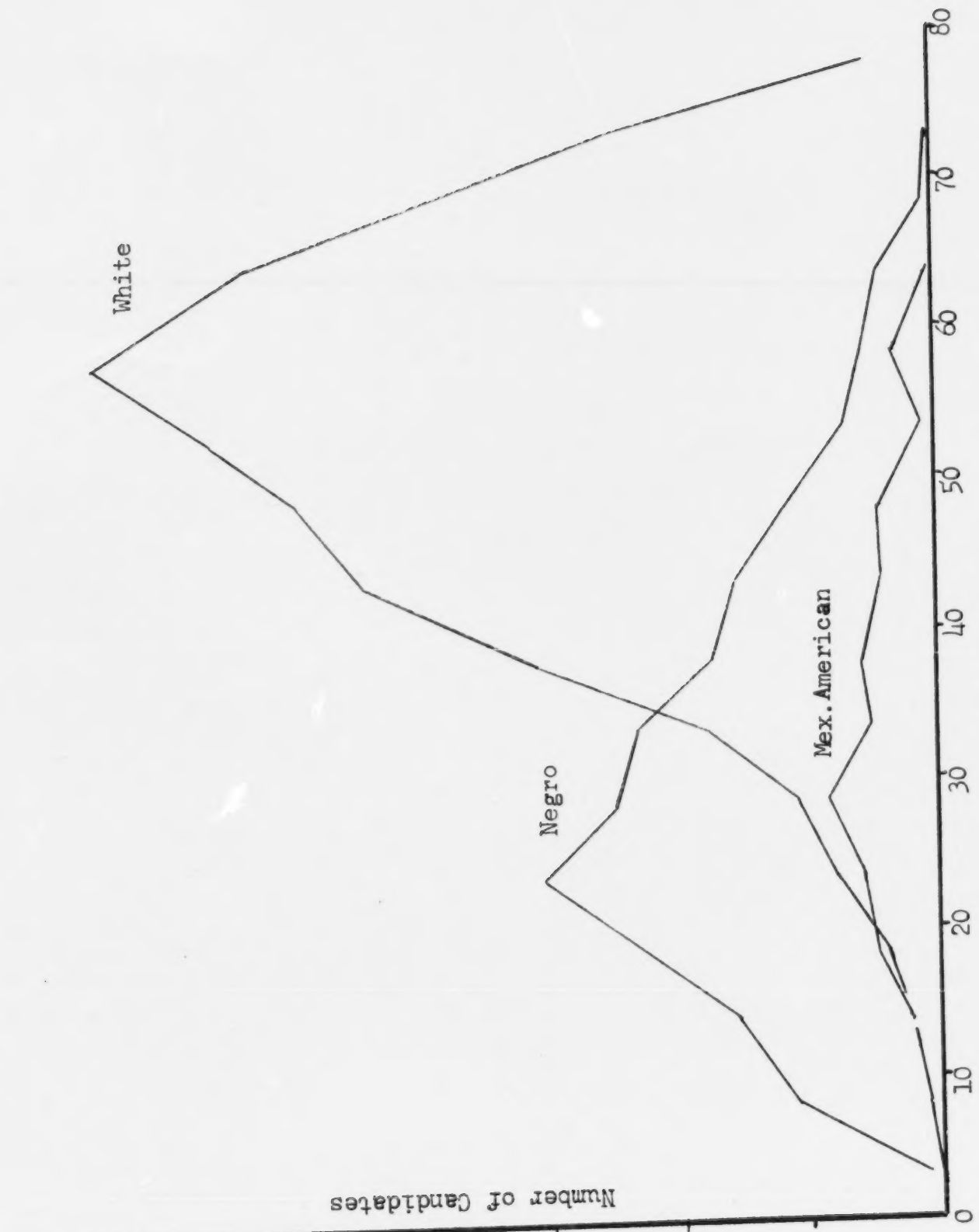
Occupation	Racial Origin			
	White	Negro	Mexican-American	Other
Account Clerk	39	3	3	1
Assistant Planner	2	1	1	0
Assistant Policewoman	38	13	0	0
Assistant Poundmaster	6	3	1	0
Auto Partsman	44	5	3	3
Budget and Research Analyst	5	0	0	1
Fireman	146	37	9	2
Gardener	15	15	0	10
Housing Inspector	15	0	0	2
Laborer	19	129	7	3
Mechanic Helper	3	29	0	1
Parking Attendant	7	11	1	0
Parking Foreman	20	0	3	2
Parking Meter Repairman	20	2	1	0
Parks Admin. Assistant	9	0	1	0
Patrolman	153	31	1	0
Police Clerk	44	1	5	2
Police Tech. III	1	2	0	0
Programmer-Operator	4	0	0	0
Public Health Statistician	7	1	0	0
Recreation Leader	85	12	6	1
Registered Nurse	22	5	1	1
Sanitarian	2	0	0	3
Secretary	15	2	0	1
Supervising Clerk	2	1	0	0
Supervisor Field Collector	45	1	0	0
Waste Collector	57	84	50	1
TOTAL	825	388	93	34

TABLE 2

Means, Standard Deviations and N's of Applicants for Selected Occupations

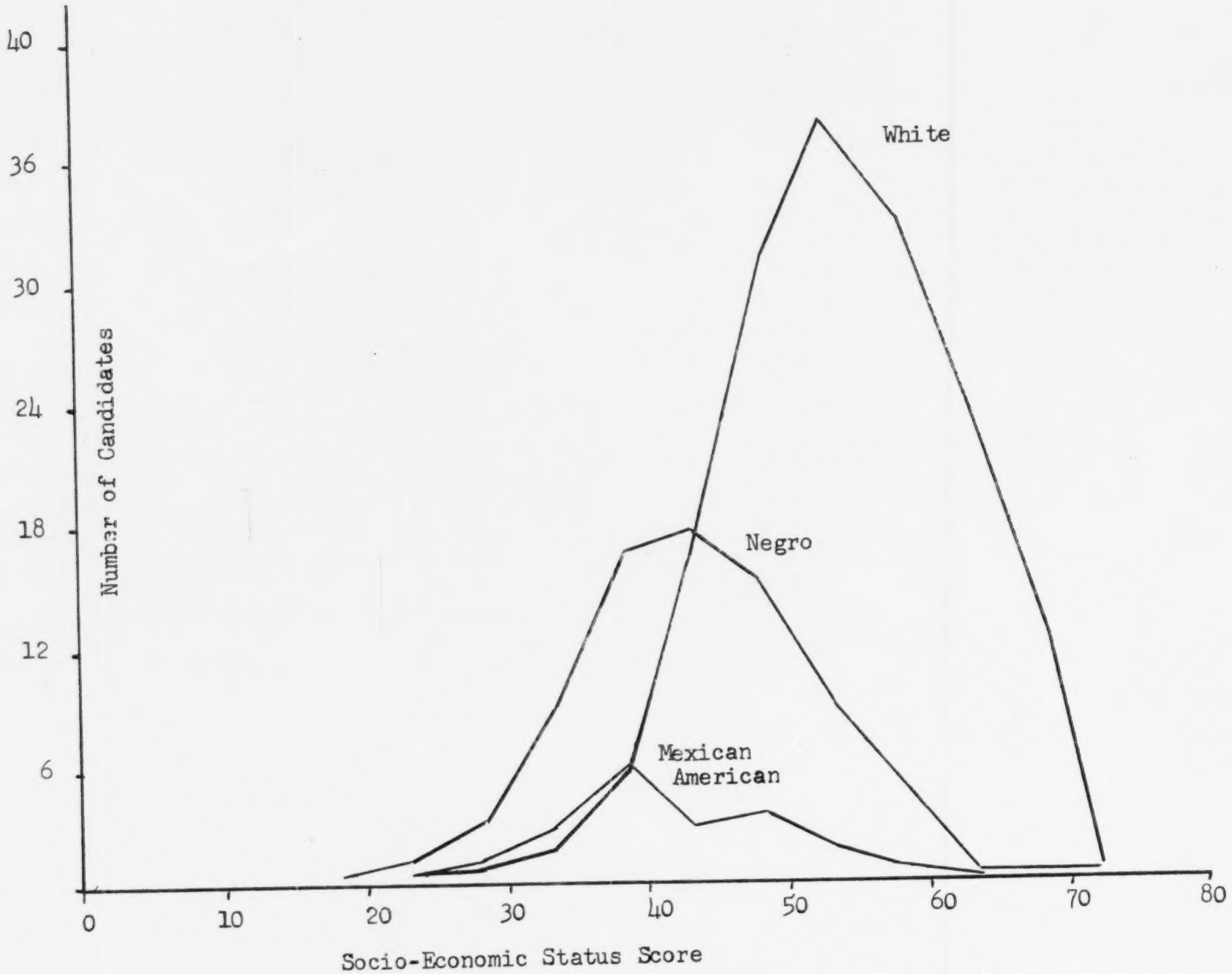
		WHITE		NEGRO		MEXICAN AMERICAN	
		OTIS	SES	OTIS	SES	OTIS	SES
Patrolman	Mean	60.6	57.8	44.9	50.7	----	----
	S.D.	9.6	6.5	12.2	5.9	----	----
	N	153	153	31	31	----	----
Fireman	Mean	55.1	51.8	40.1	46.6	----	----
	S.D.	11.0	6.0	11.7	6.2	----	----
	N	146	146	37	37	----	----
Recreation Leader	Mean	56.0	60.0	39.3	53.4	----	----
	S.D.	9.5	6.5	8.7	6.3	----	----
	N	85	85	12	12	----	----
Asst. Policewoman	Mean	53.3	54.5	46.5	50.2	----	----
	S.D.	12.6	6.4	8.9	9.4	----	----
	N	38	38	13	13	----	----
Gardener	Mean	46.3	48.2	28.7	44.5	26.5	48.1
	S.D.	17.0	9.0	8.3	6.1	9.8	4.0
	N	15	15	15	15	10	10
Laborer	Mean	42.1	45.9	23.9	39.2	----	----
	S.D.	15.4	8.6	12.0	6.9	----	----
	N	19	19	129	129	----	----
Waste Collector	Mean	37.5	43.7	25.8	41.0	29.0	37.0
	S.D.	12.3	6.3	10.3	6.8	9.8	5.9
	N	57	57	84	84	50	50
Total Group	Mean	52.4	54.2	30.2	43.2	34.1	41.1
	S.D.	13.0	8.4	14.0	8.4	11.9	8.2
	N	825	825	388	388	93	93

FIGURE 1



Frequency distribution showing the number of Otis items correctly answered by White, Negro, and Mexican-American candidate

FIGURE 2



Frequency distribution showing the Socio-Economic Status Score (SES) for White, Negro and Mexican-American candidates.

more favorable than those of the Negroes and Mexican-Americans. The white applicants also excel on test performance in comparison to the two minority groups whose test performance is substantially below the national average.

The data presented in Table 3 show the relationship among the variables included in the SES index as well as their relationship with SES and OTIS scores. Inspecting the correlations of the socio-economic variables with the OTIS (Col. 6, variables 1 through 4), it appears that the decision to combine these variables into a global SES measure giving each component equal effective weight is justified since all the correlations are moderately high. In addition, the intercorrelations of the socio-economic variables are sufficiently low, suggesting that these measures may reflect different dimensions of our socio-economic structure. The correlation of .62 between SES and OTIS is sufficiently high to warrant using SES as a covariate.

A direct comparison of the observed group means indicates that there exist very substantial differences among the racial groups on the OTIS and in socio-economic characteristics (SES). Thus, the analysis of covariance technique was used to assess the significance of group differences on test scores, using SES as a covariate. In so doing, it was anticipated that differences in the observed group means would be reduced and the magnitude of the reduction would probably warrant consideration of such social characteristics when evaluating the test performance of ethnic groups.

A summary of the analysis of covariance is presented in Table 4.

TABLE 3

Correlations, Means and Standard Deviations for the Entire Sample (N = 1340)

CORRELATION MATRIX

	Applicant's Education	Mother's Education	Father's SCI	Applicant's SCI	SES	OTIS
1	1.0000	.3002	.2591	.3655	.7041	.4982
2	.3002	1.0000	.3918	.1827	.6827	.3787
3	.2591	.3918	1.0000	.2611	.6962	.3935
4	.3655	.1827	.2611	1.0000	.6582	.4198
5	.7041	.6827	.6962	.6582	1.0000	.6151
6	.4982	.3787	.3935	.4198	.6151	1.0000
Mean	12.9	10.6	5.4	3.0	50.0	44.3
S.D.	2.0	3.1	3.1	1.7	10.0	16.9

The results of the Gulliksen-Wilks[2] tests show that the assumptions of homogeneity of within-class regression and standard error of estimate underlying the analysis of covariance model were met. But the hypothesis of equal intercepts was rejected indicating that significant group differences exist in test performance among the racial groups, even though the socio-economic status was held constant.

Figure 3 shows the regression lines for the three racial groups. Although the slopes are not entirely parallel, the divergence is not significant ($p > .10$). The outstanding difference occurs in the intercepts of the racial groups. These results indicate that on the average there is 14.7 raw score points difference between the test performance

---

[2] Gulliksen, H., & Wilks, S.S., "Regression tests for several samples." Psychometrika, 1950, 15, 91-114.

TABLE 4

Summary of the Analysis of Covariance

Analysis of Variance Table

Source	SS	DF	MS	F	P
Racial Groups	140941.118	2	70470.559	400.420	<.001
Error	228964.863	1301	175.991		
Total	369905.981	1303			

Gulliksen-Wilks Tests

Test of Equality	Chi-squared	DF	P
Errors of Estimate	1.7871	2	>.50
Slopes	3.8083	2	>.10
Intercepts	265.9396	2	<.001

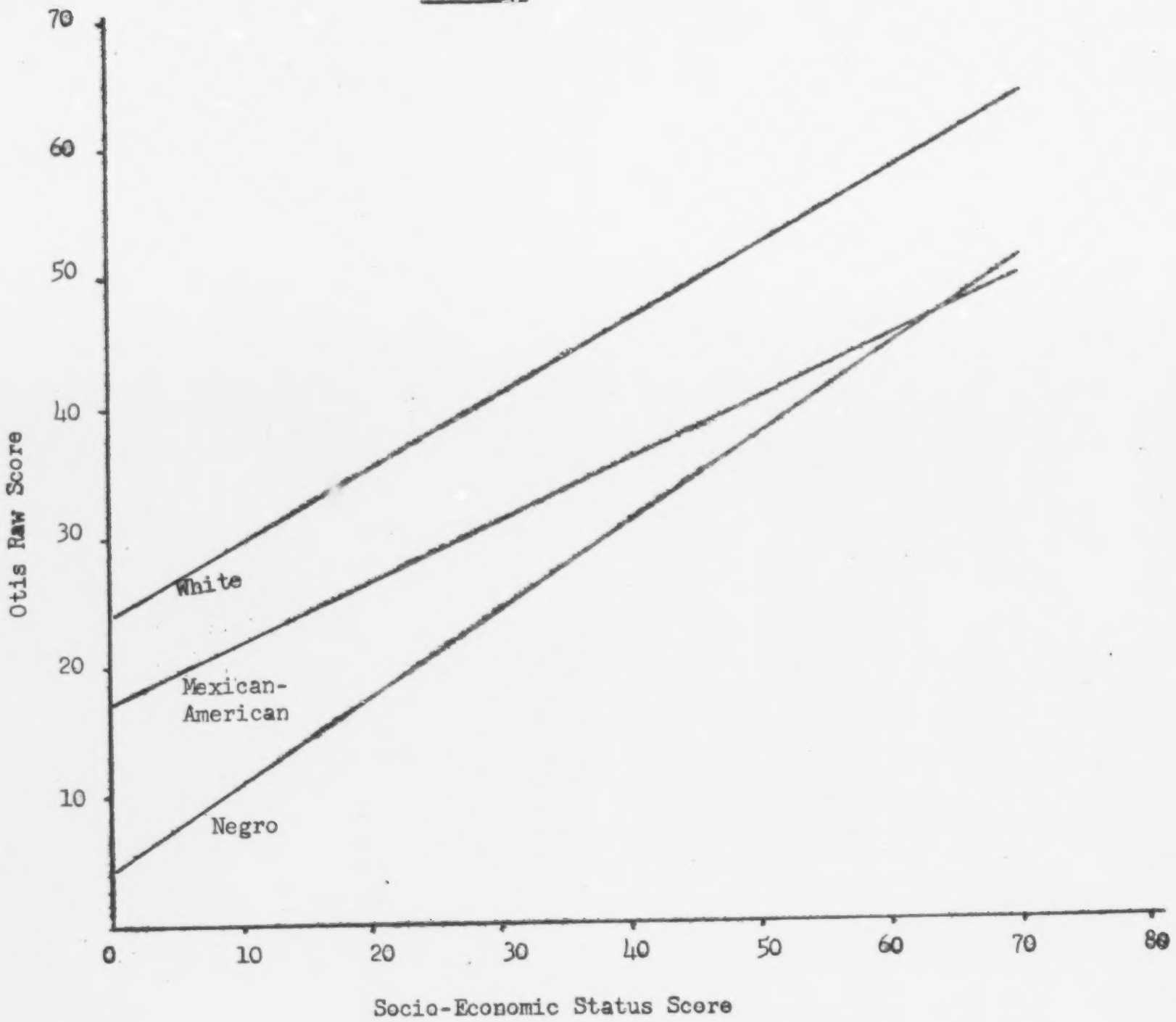
Analysis of Covariance Table

Source	SS	DF	MS	F	P
Racial Groups	42365.179	2	21182.590	147.047	.001
Error	187268.972	1300	144.053		
Total	229634.150	1302			

Adjusted Treatment Means

	White	Negro	Mexican-American	Total
SES Mean	54.218	43.162	41.054	50.009
Otis Mean	52.393	30.172	34.076	44.503
Adjusted Otis Mean	49.541	34.814	40.147	41.500

FIGURE 3



Regression lines for Socio-Economic Status (SES) and Otis raw score for White, Negro and Mexican-American candidates.

of white applicants and Negro applicants, and 9.4 raw score points between white applicants and Mexican-American applicants when differences in SES are held constant. The difference between Negro and Mexican-American applicants is 5.3 raw score points.

It should be noted that the racial distribution varies considerably among the applicants for the different job categories. It may be that there is some sort of "occupation SES race" interaction which causes sample stratification to differ widely from that found in the population. Perhaps if it were possible to obtain a sample containing equal numbers of white, Negro and Mexican-American candidates applying for each position, these differences would be reduced.

The analysis of variance design used in this study allowed the testing of three hypotheses:

(1) That there is no difference in mean item scores for the two racial groups.

(2) That there is no difference in the mean scores of different items.

(3) That there is no interaction between items and racial origin.

Table 5 contains a summary of the analysis of variance for the two racial groups. Since the item variance is a direct function of the number of applicants answering the item and most applicants attempted the first 56 items, it was decided to limit the analysis to these items to meet the assumption of homogeneity of variance.

From the data in Table 5 it is noted that the three hypotheses were rejected, implying significant differences among the item means, between the racial groups and a significant interaction between items

TABLE 5

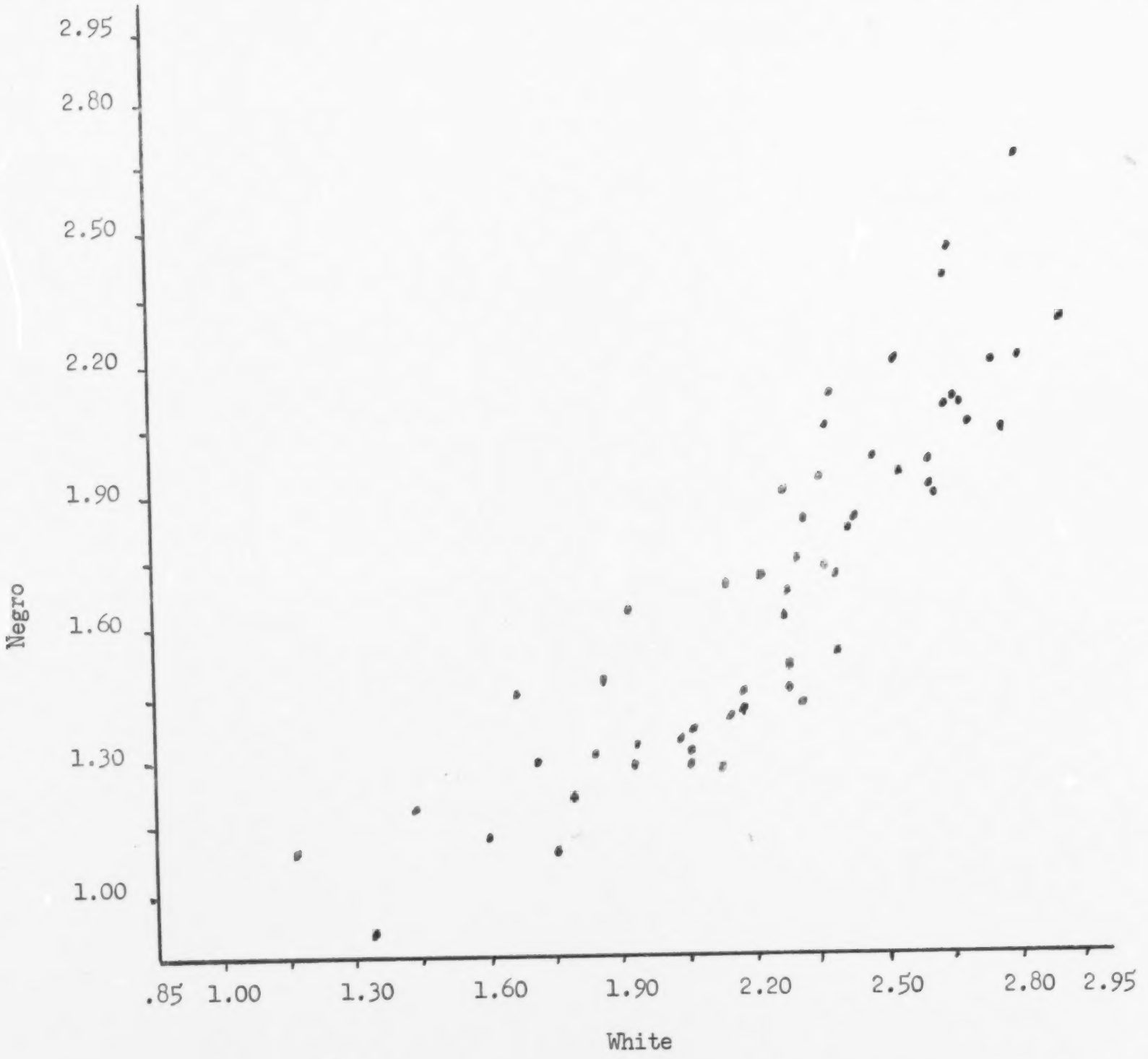
Analysis of Variance Table for the First 56 OTIS Items

Source	SS	DF	MS	F	P
Between Samples	18.270	3			
Groups	18.247	1	18.247	1575.104	<.001
Samples Within Groups	.023	2	.011		
Within Samples	34.204	220			
Items	31.699	55	.576	127.64	<.001
Groups x Items	2.005	55	.036	8.046	<.01
Items x Samples within groups	.498	110	.004		

and race. It should also be noted that most of the variance is due to differences between racial groups and the least variance is accounted for by the item x race interaction. Because of the lack of a practically significant amount of item x race interaction and because every item was more difficult for Negroes than for whites, no attempt was made to identify those items which contributed most to the item x race interaction. It was felt that a more meaningful way of looking at the results of this analysis was to present the items mean difficulty for the two racial groups in the form of scatter plots.

Figure 4 is a bivariate plot of the item mean difficulty for Negro and white applicants. From this plot, it can be seen that every item on the average is easier for the white applicants. The concentration of points below the 45 degree line indicates a systematic deviation from the expected straight line. Therefore, it must be concluded that in this sample, at least, Negro applicants did not perform as well as their white counterparts. Part of this difference in item difficulty

FIGURE 4



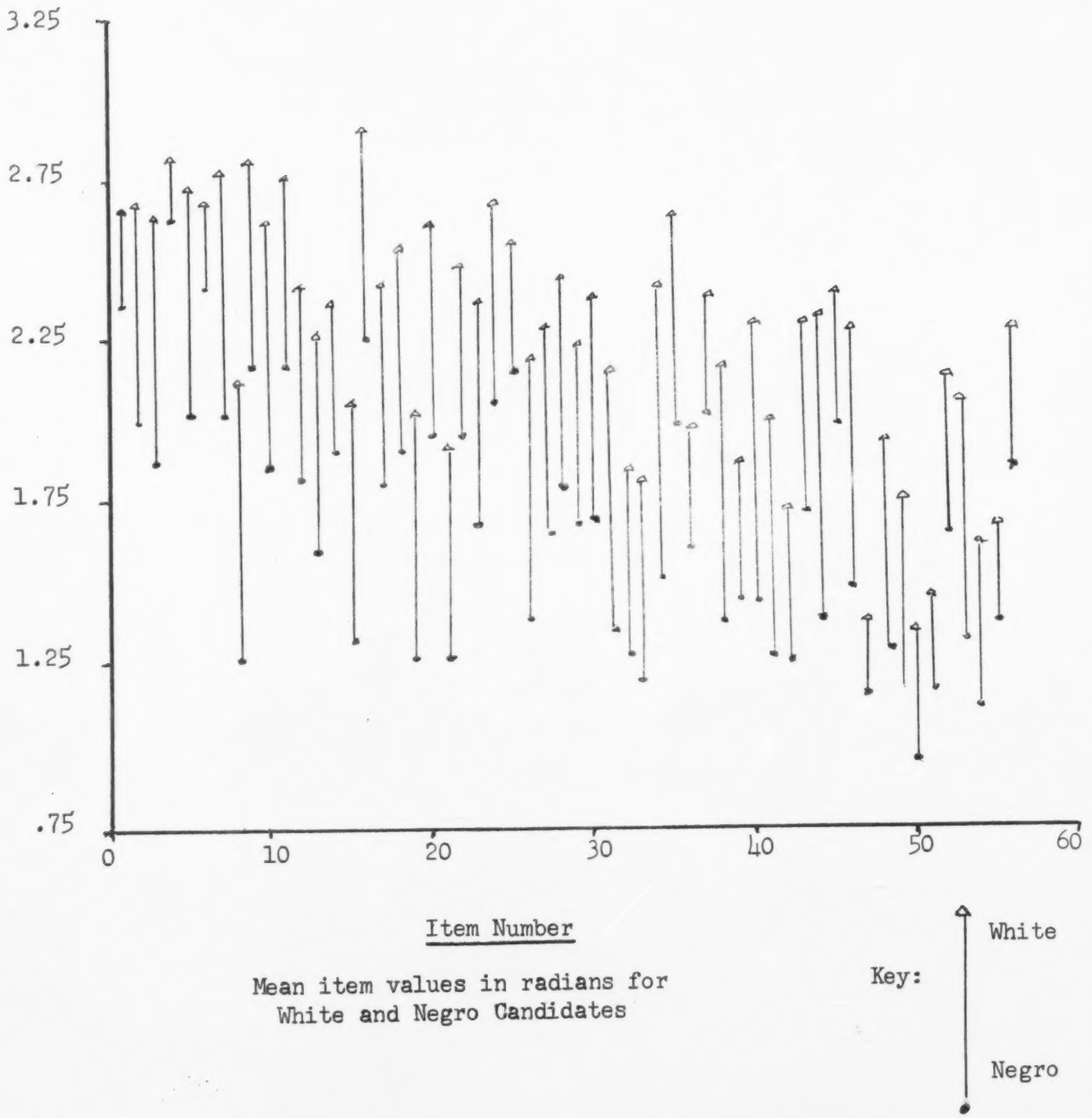
Bivariate plot of the transformed item-difficulty values for White and Negro candidates.

may be due to the differential representation of whites and Negroes among the various occupations.

Figure 5 shows each of the mean item difficulties for the two racial groups. As expected these means tend to be ordered from high (indicating relatively easy items) to low (indicating relatively difficult items), implying that the items are ordered in the test on the basis of their difficulty. More important, however, is the fact that the white applicants found each of the 56 OTIS items easier than did the Negro applicants. Again these results should be interpreted with due consideration of the differential racial representation among the various occupations.

It should be noted that the differences between item difficulties for the whites and Negroes is generally quite large; items 1, 4, 6, and 47 are the outstanding exceptions. It becomes apparent from Figure 5 that the interaction found between items and race is trivial compared to the average difference in item difficulty for the two racial groups.

FIGURE 5



Mean item values in radians for
White and Negro Candidates

CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

The data reported here seem to yield several clear, and not very surprising, conclusions.

1. In these municipal civil service examination procedures, minority group members (specifically Negro and Mexican-Americans) find the OTIS Mental Ability Test more difficult than the whites with whom they compete.
2. The minority group applicants occupy a substantially lower position in the socio-economic hierarchy than their white counterparts.
3. Lower total scores on the OTIS test are <u>not</u> attributable to a few especially "loaded" items. Though some items are relatively more difficult than others, <u>every</u> item on the test was more difficult for the minority group members.
4. A socio-economic index based on occupational status and education does not totally explain the differences in test scores among ethnic groups. However, a substantial correlation between OTIS score and socio-economic status was found. (.62)

The implications of these results for improvement of the personnel selection process are obviously limited. It was clear at the beginning of this study that the question of test validity is crucial. These findings suggest that the traditional methods of determining test validity are not appropriate in this setting.

With knowledge of the striking difference in test performance and

socio-economic status of ethnic groups, it cannot be assumed that, (a) tests such as the OTIS are equally valid throughout their score range, nor (b) that such tests are equally valid for different racial groups at any given score point. Since test scores are frequently used as cutting points at some stage in the selection procedure, the maximally-valid cutting point for particular jobs -- and specified groups should be determined. It seems quite likely that differing racial, socio-economic, and educational characteristics will be associated with different, maximally-valid, cutting points. This does not suggest lowering standards or giving some groups special advantages in the personnel selection process. It is simply a matter of using the best available techniques to achieve increased efficiency in personnel selection.

The application of decision theory to personnel selection has been discussed in Psychological Tests and Personnel Decisions by Cronbach and Gleser (Univ. of Ill., Press, 1965). Two passages from this book seem appropriate:

> "A tremendous amount of knowledge has been developed around the general topic of decision problems. Since the tester is concerned with decision making, it is reasonable to expect significant understandings to result from restating testing problems in such a way that this knowledge can be brought to bear." (p. 4)

> "The conceptual tools of decision theory make the tester aware of problems that have hitherto been minimized or overlooked. ---restatement of testing problems in these new terms will assist all users of tests to understand what assumptions they are making and how adequate their procedures are." (p. 5)

After cutting points have been established, it is common practice to rank (on the basis of test score) those applicants who survive the cutting point decision. The validity and reliability of such rankings

are questionable. If the utility of a test has been exhausted in its use as a cutting point criterion, it should not be used in the ranking process.

This study has been concerned with racial factors in personnel selection procedures. It should be recognized that the criterion performance may often be racially biased. For example, one may select candidates who will succeed in a training program. If the training program puts the minority group trainee at a disadvantage, "successful" selection will predict his poor performance.

The efforts of the Commission's TACT committee through such publications as the Guidelines for Testing and Selecting Minority Job Applicants are one way to reduce racial bias in employment practices. Efforts to develop better, and therefore more equitable, techniques for personnel selection are another.